Proposal of a Visualization System for a Hierarchical Clustering Algorithm: The Visualize Proximity Matrix

Sulaiman Abdullah Alateyah[1]

[1] Department Of Computer Science, College Of Science and Arts, Qassim University, Unaizah, Saudi Arabia;

Correspondence: Sulaiman Abdullah Alateyah Department Of Computer Science, College Of Science and Arts, Qassim University, Unaizah, Saudi Arabia. E-mail: salateyah@qu.edu.sa

Abstract

Visual data mining is a new and effective strategy for dealing with the growing phenomenon of information overload. There is an urgent need for effective visual data mining because the growth of data sets from different domains and sources has made exploring, managing and analyzing vast volumes of data increasingly difficult. Identifying location-based trends and anomalies in data from a numeric output is challenging. The outputs of data mining procedures are often quite difficult to interpret. There is no human input to data exploration or direct involvement in the data mining process. Thus, it is difficult to analyze, explore, understand and view the data. In data mining, massive data sets are clustered using an iterative process that includes human input but data mining algorithms can miss some knowledge and observations. In addition, many users cannot take the appropriate action to address problems or opportunities because the huge amounts of data prevent them from staying aware of what is happening in and around their environments. Visual data mining can help in dealing with information overload and it is effective in analyzing large complex data sets. Visual data mining assists users in gaining a deep visual understanding of data. In the information age, users need to study and observe vast amounts of data to acquire important knowledge, and thus the need for visual and interactive analytical tools is particularly pressing. Visual cluster analysis has long utilized shaded similarity matrices, and this study investigated how they can be used in clustering visualization. We focused on proposing an agglomerative hierarchical clustering method. Ensemble cluster visualization is also presented for handling large data sets. This study proposes the adoption of a shaded similarity matrix to visually cluster knowledge discovered using data mining. Using the technology acceptance model as the measuring tool, we questioned respondents to evaluate the visualization prototype. The findings demonstrated that the visualization was effective and easy to use, and satisfied users.

Keywords: Data mining, Visual data mining, Pattern, Algorithm, Exploring

1. Introduction

In this information era, data are collected from many sites and sources such as loop detectors, marketing, technology, medical and banking. Data mining algorithms are applied to detect hidden knowledge or patterns. There are many data mining algorithms that can be adopted for analysis; for example, clustering, decision tree, genetic algorithm and neural networks. Data mining is the result of a long process of research, development and evaluation (Card, 1999; Guo et al., 2020; Spence, 2001; Ware, 2012).

In their work, (Pampalk et al., 2003; Schneiderman & Plaisant, 2005; Wickens & Hollands, 2000) stated that complicated data mining techniques were necessary for the study of complex and heterogeneous data. With new visualization paradigms, many analysis approaches can be applied which benefit from visual data processing. Visualization is a natural way to combine several data sources, has been applied in many different fields and has been confirmed to be reliable and effective. Although visual approaches cannot replace data mining algorithms, it is useful to combine visualization techniques and data mining algorithms in data exploration processes. The main purpose of visualization is to convert data into an appropriate representation or visual form. Then, users can use their recognition skills to interpret, understand, observe, analyze and query the data efficiently. Users should be directly involved in the analysis of the data in order to maximize the usefulness of the visualization tool. They should also be able to dynamically explore the visual representation of the data in order to comprehend them more quickly and easily.

As a result, visual presentation can be extremely effective to highlight patterns, outliers, clusters and data gaps.

This research also aimed to visualize the result extracted by a data mining algorithm called knowledge which enables users to interact with every step of the data mining process, making it easier to interpret and view the data. According to (Keim, 2002), Visual Data Mining (VDM) is a new and effective strategy for dealing with increasing information overload. Never in history have there been as many data points produced as there are now. Effective visual data mining, data mining and visualization are now necessary because of the difficulty of exploring, managing and interpreting the huge volumes of data in the form of ever-expanding data sets from many fields and sources (Meyer & Cook, 2000). Visual data mining is an efficient way to process huge, complicated data sets and can help with information overload. To fully grasp data visually, visual data mining is necessary (Feng et al., 2021; Mendoza-Silva et al., 2021). This study proposes a visual cluster approach to visualize the knowledge extracted by a data mining algorithm based on a tree strategy for monitoring the data, involving the user in the data discovery process and allowing the user to analyze and observe large amounts of data in order to extract valuable knowledge.

2. Literature Review

In order to extract hidden rules, expressions and meaningful patterns from these big data, scalable and reliable analytical algorithms must be developed. Huge amounts of data are amassing in numerous fields, including in scientific and engineering databases.

Visual data mining is an emerging interdisciplinary science aiming to develop automatic or semiautomatic techniques which can discover the knowledge hidden in these databases, to make decision-making processes faster and more efficient. Hence, utilization of data mining in medical, education, finance, engineering, marketing and telecommunication industries has dramatically increased in recent years. Incorporating data mining algorithm and visualization methods is potentially effective, as revealed by successful visual data mining tools such as generative topographic mapping and the self-organizing map. However, a significant amount of integration work remains to be done in order to benefit from advanced results from both domains.

There are various methods for visualizing data; for example, x–y plots, line plots and histograms. These procedures are valuable for data exploration but are generally restricted to small and low-dimensional data sets. In recent decades, a large number of novel data representation methods have been developed, enabling visualization of multidimensional data sets (Card, 1999; Guo et al., 2020; Ran et al., 2023; Spence, 2001; Ware, 2012).

The strength of visualization represents the capacity for discovery (Schneiderman & Plaisant, 2005; Wickens & Hollands, 2000) Implementing visualization tools to examine and comprehend high-dimensional information is currently proving to be an effective method of combining intelligence with the enormous capabilities of the processing power currently available (Pampalk et al., 2003; Yang & Hussain, 2023).

The key benefit is making use of the human visual system to assist the data mining process. This is achieved through the creation of visualizations of the data which allow users to identify features within the data which would not otherwise be apparent (Feng et al., 2021; Keim, 2002; Mendoza-Silva et al., 2021; Meyer & Cook, 2000). According to (Zhang, 2008), the human visual system comprises the brain and the eyes. The eyes can be regarded as a strong and highly parallel processing and reasoning engine.

(Bhadran et al., 2008), explained that visualization techniques are widely used in exploring, understanding, summarizing, interpreting, observing and analyzing large amounts of data (Ankerst et al., 2000; Grinstein & Wierse, 2002; Morrison et al., 2002; Shneiderman, 2001). Many different visualization approaches, including geometric, icon-based, pixel-oriented, hierarchical and graph-based methods, have been created to map multidimensional data sets to two- or three-dimensional space.

3. Proposed Visualizing Proximity Matrix

The shaded similarity matrix is described in this section. For the past 40 years, visual cluster analysis has primarily used shaded similarity matrices. In (Gale et al., 1984; Ling, 1973; Ran et al., 2023), the authors provide a full summary of the early work, whereas (Biedl et al., 2001; Wishart, 1999) include some recent work that is pertinent. Greater likeness is shown by darker shading, whereas lower resemblance is represented by lighter shading. Dark and light cells may be spread throughout the matrix at first, so the rows and columns are restructured so that similar things are placed adjacent to one another to show possible groupings. If there are "real" clusters in the data, they should show up as symmetrical black squares on the diagonal (Cao et al., 2023; Gale et al., 1984). This tutorial will explain how a shaded similarity matrix is constructed and how it looks using an example. The data for this example are taken from the literature data set (Merz & Murphy, 1996; Zhu et al., 2018; Zidan et al., 2020).

The dendrogram tree method is used in hierarchical cluster analysis to visualize how the cluster is merged. The visualizing proximity matrix proposed in this study shows the cluster in a contrasting color and also shows the distance between the merged clusters. The following snapshots explain the agglomerative hierarchical clustering and visualizing proximity matrix. Fig 1 shows the main page of the proposed agglomerative hierarchical clustering.
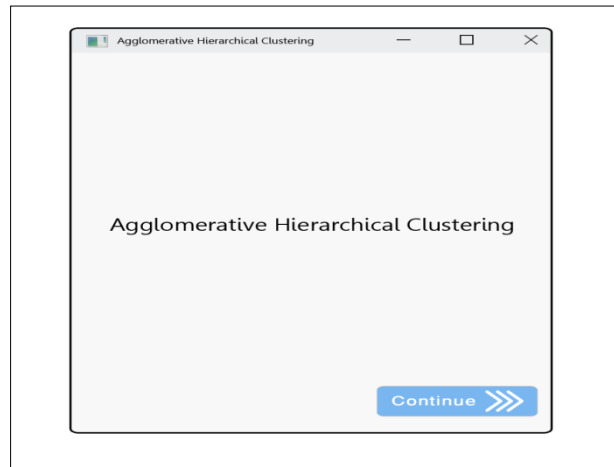

Figure 1. Main Page

The similarity measure dialog box is shown in Fig 2. This specifies the distance measure and the clustering method. The first step is to select the similarity distance measure. For interval data, the most common measure is Euclidian distance.
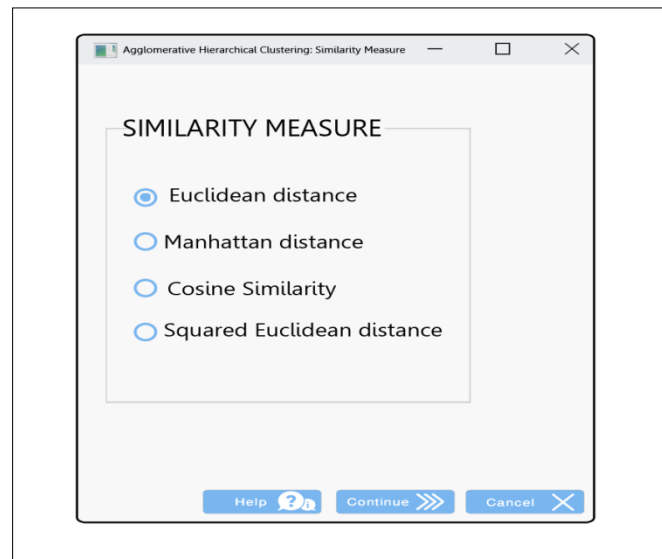

Figure 2. Similarity measure.

Complete, single and average hierarchical clustering methods are the linkage methods used to calculate the distance between data points. These are all based on Euclidean distance but the main difference between them is the selection of the data points that are considered as the final criterion on which the similarity or distance depends, as shown in Fig 3.
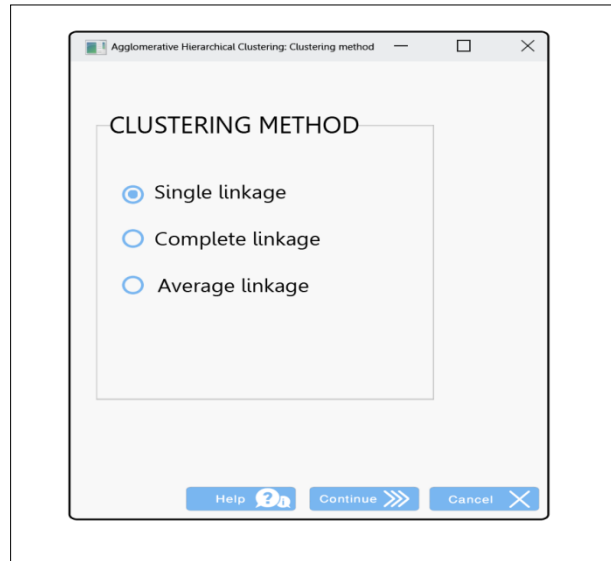
Figure 3. Clustering method.

Fig 4 shows the visualization dialog box. The Visualize Proximity Matrix and the dendrogram will graphically show how the clusters are merged and enable identification of the appropriate number of clusters. Then, the user chooses a specified range of clusters or all clusters, as shown in Fig 5.
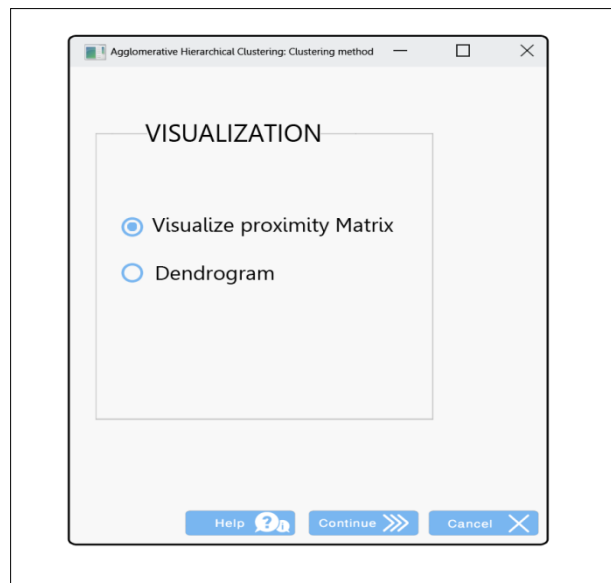
Figure 4. Visualization.

Figure 5. Specified Range Of Clusters Or All Clusters.

Using single linkage clustering, Fig 6 shows that the nearest pair is "D" and "F", at a distance of 0.50. The D and F are merged together into one cluster called "D, F". Fig 7 shows that the nearest pair is "A" and "B", at a distance of 0.71. The A and B are merged together into one cluster called "A, B". Fig 8 shows that the nearest pair is "E" and "D,F", at a distance of 1.00. The D, F and E are merged together into one cluster called "D, F, E". Fig 9 shows that the nearest pair is "C" and "D, F, E", at a distance of 1.41. The D, F, E and C are merged together into one cluster called "D, F,E,C". Fig 10 shows that the nearest pair is "A,B" and "D, F, E,C", at a distance of 2.50. Finally, these clusters are merged together into a single cluster called "A,B, D, F,E,C".
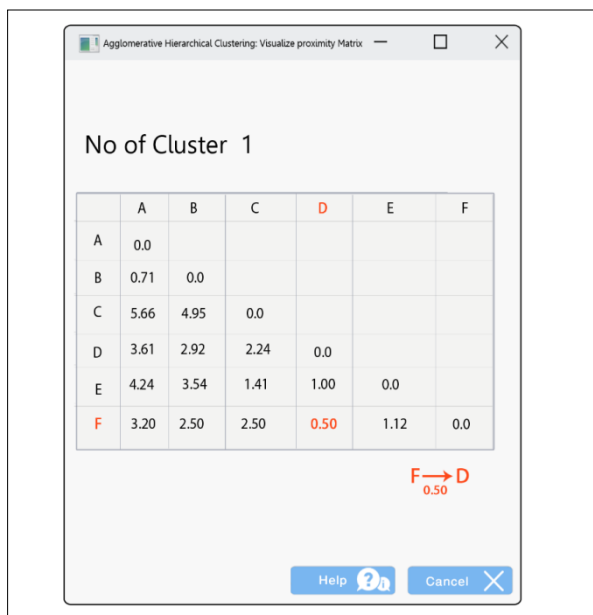

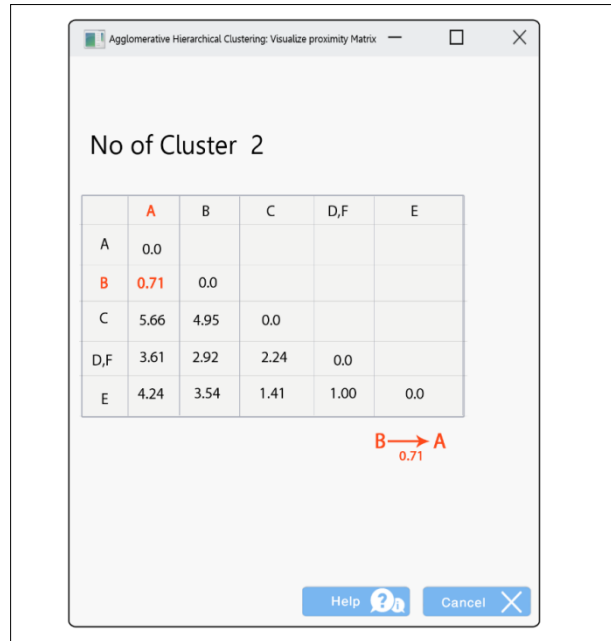Figure 6. Visualize Proximity Matrix for Cluster 1.

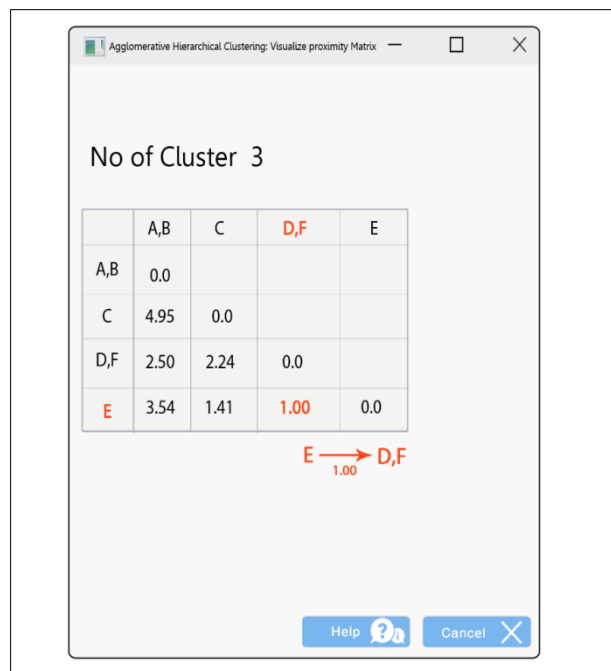Figure 7. Visualize Proximity Matrix for Cluster 2.

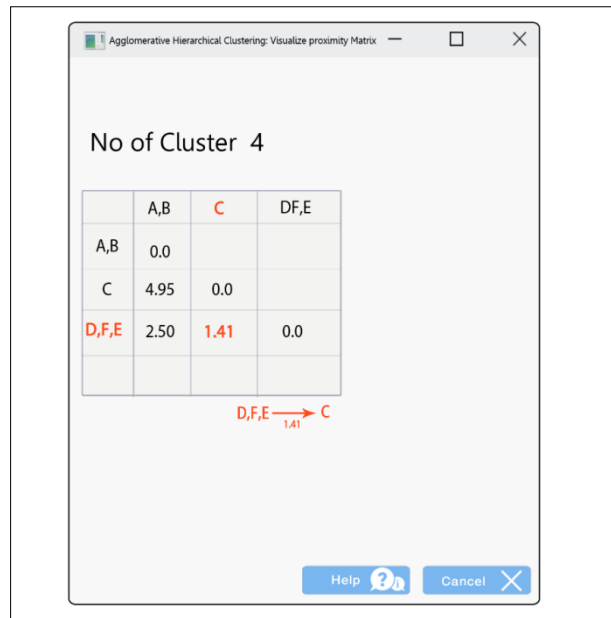Figure 8. Visualize Proximity Matrix for Cluster 3.

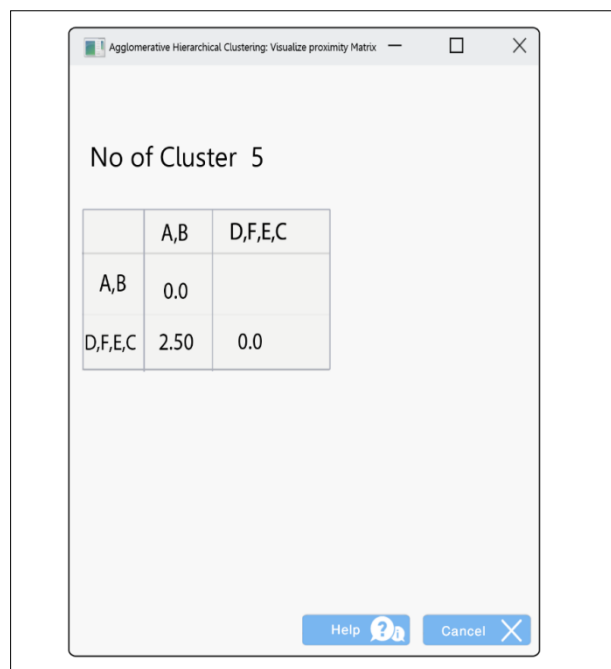Figure 9. Visualize Proximity Matrix for Cluster 4.

Figure 10. Visualize Proximity Matrix for Cluster 5.

4. Methodology and Hypothesis Development

Examination of the proposed proximity matrix quantifies its viability and determines whether the targeted interest group has reached the indicated locations. Evaluation should cover both the incorporation of the visualization with data mining algorithms, as well as the usability and usefulness of the visualization element for controlling the data. It should include the role of the user in the data exploration process, and whether it enables users to examine many facts to obtain useful information. According to (Kanaujiya, 2008), a visual data mining prototype must be syntactically simple to be useful. To be simple to learn, it needs to be easy to extract and interpret knowledge using intuitive and user-friendly tools. To be simple to apply, it needs to allow efficient communication between humans and data. Questionnaires have long been used to evaluate software systems and user interfaces (Root & Draper, 1983). The biggest single advantage of using questionnaires in evaluation of an interactive prototype is that they provide data on prototype acceptance from the user point of view.

(Yamazawa et al., 2008) visualized the drags based on chemical structures, after dividing the drags using the hierarchical clustering method. They evaluated user experiments, discussed the effectiveness of the presented technique using the level of detail (LOD) control technique. Eleven examinees were instructed to operate and explore the user interface in the LOD for a given time, and then rate it. All the participants in the assessment were either postgraduate or undergraduate information science students.
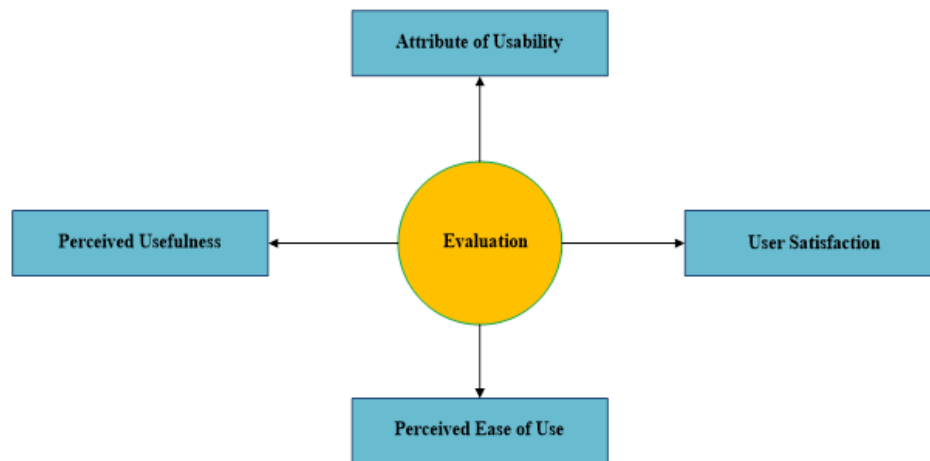

Figure 11. Evaluation Dimensions.

The sampling procedure that was adopted in this study for data collection was a target sampling method using a questionnaire survey issued to 38 respondents. The sample consisted of 23 males and 15 females and 27 were in the age group 25–35. Generally, the respondents had seen and watched the visualization prototype of the bidirectional agglomerative hierarchical clustering algorithm. The demographic profile of the respondents is presented in Table 1.

Table 1. Demographic Profile

Variable		Frequency	Percentage
Gender	*Female*	23	61%
	Male	15	39%
Age	25-35 years	27	71%
	36-45 years	6	16%
	46 >	5	13%

4.1 Dimensions

4.1.1 Perceived Usefulness

This defines the degree to which users believe that observation and exploration of knowledge will be improved by use of the visualization prototype. Table 2 shows the items used to measure the dimension of perceived usefulness. The first dimension examines HYPOTHESIS 1: Through visualization it is possible to observe and explore knowledge within the data that has been missed by data mining algorithms, and working with the visualization output is easier than working with a numeric output.

4.1.2 Perceived Ease of Use

This refers to the degree to which users believe that using the visualization system is effortless and that interaction with the visualization system is clear and understandable. Table 2 shows the items used to measure the dimension of perceived ease of use. The second dimension examines HYPOTHESIS 2: Interaction with visualization is clear, understandable and effortless. It is easy to become skilled at using visualization to explore and observe knowledge.

4.1.3 User Satisfaction

This refers to what users expect from the system, and therefore is a personal assessment about what the system should do for the end-user. Table 2 shows the items used for measurement of the user satisfaction dimension. The third dimension examines HYPOTHESIS 3: It is easy to be aware of what is happening in and around their environments from the huge amounts of data and to draw out insights by facilitating commentary or discussion regarding the experience.

4.1.4 Attribute of Usability

This is the area of human–computer interaction (HCI) with regard to the proposed visualization system. It attempts to bridge the gap between the goals of the user and the system. Table 2 shows the items used to measure the user attribute of usability. The fourth dimension examines HYPOTHESIS 4: Does the visualization introduce human issues into the design and devise practical techniques for the observation of human behavior and performance?

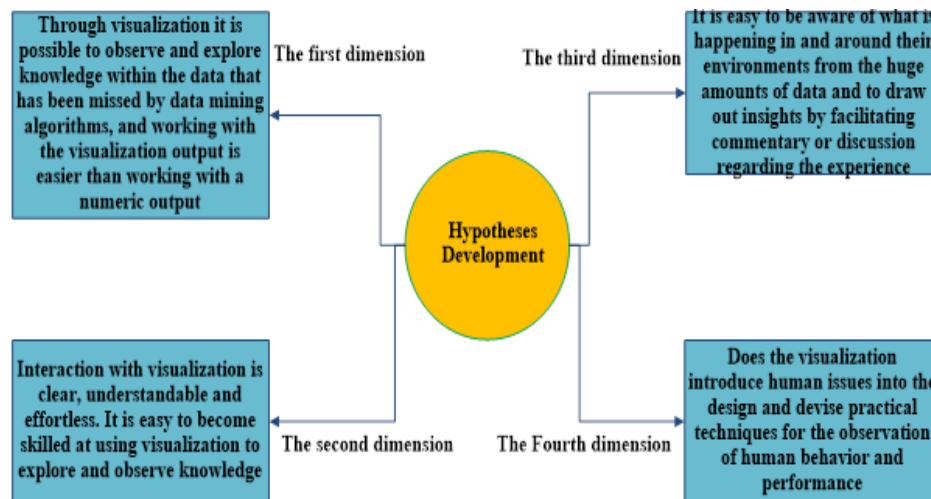The hypothesis development and the four dimensions are shown in Fig 12.

Figure 12. Hypothesis Development

5. Results and Discussion

The user testing of the visualization tool and the Visualize Proximity Matrix examined the effectiveness and usefulness of visualization tool. We asked 38 respondents to use and explore the visualization tool for several minutes and then evaluate it. All the respondents had knowledge of or connection with the fields of computer science and information technology. The descriptive statistics for the main variables in Table I revealed that all dimensions were scored higher than the midpoints of their respective scales. This shows that respondents were generally optimistic about the four dimensions of using the visualization tool to extract and view knowledge that has been missed by data mining algorithms. Additionally, users and data miners were able to be involved and interact in the data processes by exploiting the power of the human brain and visual abilities to analyze and explore data. Therefore, the visualization prototype shows promise as a useful tool to solve the problem statements for this research:

- Human beings are not included in data exploration processes.
- Some knowledge and observations are missed by data mining algorithms.

The respondents proved that working with visualization usually assists in the rapid discovery of data and understanding its structure. It was also proved that it is easier to work with a visualization output than with a numeric output, and that it is effective in enabling the users and data miners to identify and view hidden patterns and rules in data that might have been missed by the data mining algorithms.

The findings of the evaluation identify some areas of the teaching materials that require clarification. The visualization tool needs to be suitable for use by beginners and both novice and experienced users must be able to access comprehensive instructional information. Finally, based on self-reporting, some people are highly visual whereas others are not. Table 2 shows the responses to the survey concerning the four dimensions and Fig 13 shows the Hypotheses Testing for all questions.

Table 2. Dimension Items.

| Dimensions Items	Responses /Mean
First dimension to examine the HYPOTHESIS 1: "Through visualization it is possible to observe and explore knowledge within the data that has been missed by data mining algorithms, and working with the visualization output is easier than working with a numeric output"	3.96%
PU1: Using visualization tool can help the user explore, observe the knowledge easier	3.8%
PU2: Proposed Visualization tool will enable the user to get information of data quickly	3.4%
PU3: Working with visualization output is easier than working with numeric output	4.2%
PU4: By using Visualization tool humans might catch and observe hidden patterns and rules in data	4.3%
PU5: By using visualize proximity Matrix is directly involve and interactive in the data processes by exploitation the power of the human sight and brain for analyzing and exploring data	4.1%
Second dimension to examine the HYPOTHESIS 2: "Interaction with visualization is clear, understandable and effortless. It is easy to become skilled at using visualization to explore and observe knowledge "	3.54%
EU1: Learning to operate Visualization tool would be easy for me	3.1%
EU2: I find it easy get Visualization tool to do what I want it to do	3.4%
EU3: My interaction with Visualization tool would be clear and understandable	4.2%
EU4: I found Visualization tool flexible to interact with	4.1%
EU5: It is easy for me to become skillful at using Visualization tool	2.9%
Third dimension to examine the HYPOTHESIS 3: " It is easy to be aware of what is happening in and around their environments from the huge amounts of data and to draw out insights by facilitating commentary or discussion regarding the experience"	3.38%
US1: I completely satisfied in using the Visualization tool	3.6%
US2: I feel very confident in using the Visualize proximity Matrix	3.1%
US3: I found it easy to explore the data by using Visualize proximity Matrix	2.8%
US4: I can accomplish the task quickly using this procedure	3.2%
US5: I believe that from using Visualize proximity Matrix it easy to stay aware of what is happening in and around their environments from the huge amounts of data	4.2%
Fourth dimension to examine the HYPOTHESIS 4: "Does the visualization introduce human issues into the design and devise practical techniques for the observation of human behavior and performance "	3.56%
AU1: It easy to interact with Visualization tool by using Visualize proximity Matrix	4.4%
AU2: The procedure through Visualize proximity Matrix is clear	4.1%
AU3: I found the use of Visualize proximity Matrix is suitable for each community groups.	2.6%
AU4: I found the various functions in this Visualization tool were well integrated	3.5%
AU5: I think that i would like to use this Visualization tool always	3.2%
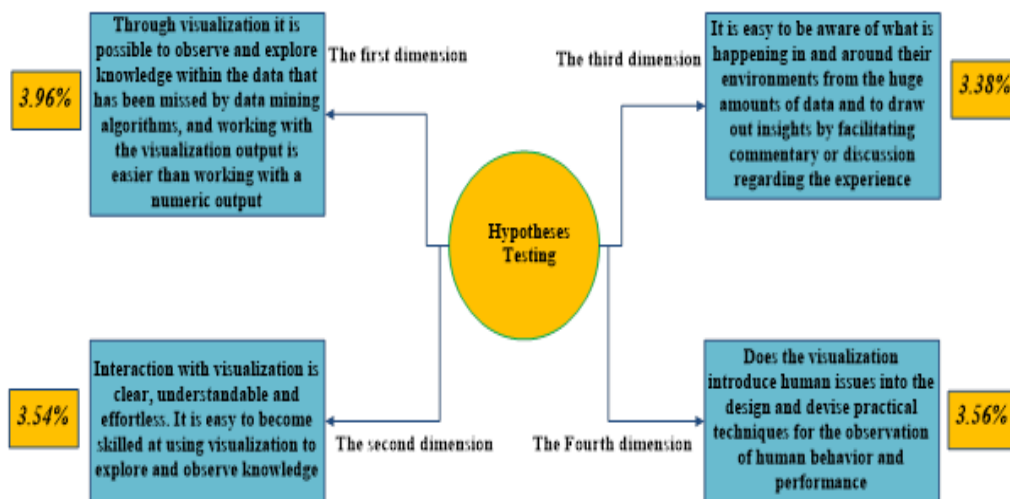
Figure 13. Hypotheses Testing.

6. Conclusion

Data mining processes involve multiple stages such as target data, data integration, data cleaning, data selection, data transformation and the output of the data mining algorithm. The objective of this paper is to propose a visualization tool, the Visualize Proximity Matrix, which will assist in knowledge discovery, understanding the

structure of the data and handling large amounts of data through the provision of an effective exploratory visualization tool. Interactive analysis tools reduce the gap between the human being and the flood of information that the human needs to search in order to extract valuable knowledge. Therefore, visual data mining has become a critical technological process, which uses visual data mining steps to avoid information overload. There is a need to be able to use cognitive abilities to transform the data into information that can eventually be used to make decisions, solve problems, improve products and increase understanding.

Visualization is a highly effective modality for understanding the structure of data and information. Based on the results of this evaluation, the visualization prototype shows promise as a tool for exploring, observing and increasing understanding of data. This evaluation enabled users and data miners to interact and be directly involved with the visualization prototype in order to easily explore and extract knowledge from the data that was missed by the data mining algorithms.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Canadian Center of Science and Education.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

References

Ankerst, M., Ester, M., & Kriegel, H. P. (2000). Towards an Effective Cooperation of the User and the Computer for Classification. *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 179-188. https://doi.org/10.1145/347090.347124

Bhadran, V., Roy, R. C., & Gopikakumari, R. (2008). Visual representation of 2-D DFT in terms of 2×2 Data: A pattern analysis. *2008 International Conference on Computing, Communication and Networking*, 1–9. https://doi.org/10.1109/ICCCNET.2008.4787762

Biedl, T., Brejová, B., Demaine, E. D., Hamel, A. M., & Vinar, T. (2001). Optimal arrangement of leaves in the tree representing hierarchical clustering of gene expression data. *University of Waterloo, Canada*.

Cao, L., Zhao, Z., & Wang, D. (2023). Clustering algorithms. In *Target Recognition and Tracking for Millimeter Wave Radar in Intelligent Transportation* (pp. 97-122). Springer. https://doi.org/10.1007/978-981-99-1533-0_5

Card, M. (1999). *Readings in information visualization: using vision to think*. Morgan Kaufmann.

Feng, X., Nguyen, K. A., & Luo, Z. (2021). A survey of deep learning approaches for WiFi-based indoor positioning. *Journal of Information and Telecommunication*, 1-54.

Gale, N., Halperin, W. C., & Costanzo, C. M. (1984). Unclassed matrix shading and optimal ordering in hierarchical cluster analysis. *Journal of Classification*, *1*(1), 75-92. https://doi.org/10.1007/BF01890117

Grinstein, U. M. F. G. G., & Wierse, A. (2002). *Information visualization in data mining and knowledge discovery*. Morgan Kaufmann.

Guo, L., Wang, F., Sang, J., Lin, X., Gong, X., & Zhang, W. (2020). Characteristics analysis of raw multi-GNSS measurement from Xiaomi Mi 8 and positioning performance improvement with L5/E5 frequency in an urban environment. *Remote Sensing*, *12*(4), 744. https://doi.org/10.3390/rs12040744

Kanaujiya, S. (2008). Visual data mining. *Proceedings of 2nd National Conference on Challenges & Opportunities in Information Technology (COIT-2008) RIMT-IET, Mandi Gobindgarh. March 29*.

Keim, D. A. (2002). Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics*, *8*(1), 1-8. https://doi.org/10.1109/2945.981847

Ling, R. L. (1973). A computer generated aid for cluster analysis. *Communications of the ACM*, *16*(6), 355-361. https://doi.org/10.1145/362248.362263

Mendoza-Silva, G. M., Costa, A. C., Torres-Sospedra, J., Painho, M., & Huerta, J. (2021). Environment-aware regression for indoor localization based on WiFi fingerprinting. *IEEE Sensors Journal*, *22*(6), 4978-4988. https://doi.org/10.1109/JSEN.2021.3073878

Merz, C. J., & Murphy, P. M. (1996). *UCI repository of machine learning databases. University of California, Department of Information and Computer Science*. Technical Report. Retrieved from http://www.ics.uci.edu/mlearn/MLRepository. html

Meyer, R. D., & Cook, D. (2000). Visualization of data. *Current Opinion in Biotechnology*, *11*(1), 89-96. https://doi.org/10.1016/S0958-1669(99)00060-9

Morrison, A., Ross, G., & Chalmers, M. (2002). Combining and comparing clustering and layout algorithms. *Tech. Rep. 148, Department of Computing Science, University of Glasgow*.

Pampalk, E., Goebl, W., & Widmer, G. (2003). Visualizing changes in the structure of data for exploratory feature selection. *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 157–166. https://doi.org/10.1145/956750.956771

Ran, X., Xi, Y., Lu, Y., Wang, X., & Lu, Z. (2023). Comprehensive survey on hierarchical clustering algorithms and the recent developments. *Artificial Intelligence Review*, *56*(8), 8219-8264. https://doi.org/10.1007/s10462-022-10366-3

Root, R. W., & Draper, S. (1983). Questionnaires as a software evaluation tool. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 83-87. https://doi.org/10.1145/800045.801586

Schneiderman, B., & Plaisant, C. (2005). *Designing the user interface: Strategies for Effective Human-Computer Interactions. The United States of America*. Pearson Education, 4th edition, Reading, MA: Addison-Wesley.

Shneiderman, B. (2001). Inventing discovery tools: Combining information visualization with data mining. *International Conference on Discovery Science*, 17-28. https://doi.org/10.1007/3-540-45650-3_4

Spence, R. (2001). *Information visualization* (Vol. 1). Springer. https://doi.org/10.1057/palgrave.ivs.9500019

Ware, C. (2012). *Information Visualization: Perception for Design, Waltham, MA*. USA: Morgan Kaufmann.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering Psychology and Human Performance. (2000)*. Upper Saddle River, NJ: Prentice Hall.

Wishart, D. (1999). Clustan Graphics3 Interactive Graphics for Cluster Analysis. In *Classification in the Information Age* (pp. 268–275). Springer. https://doi.org/10.1007/978-3-642-60187-3_27

Yamazawa, M., Itoh, T., & Yamashita, F. (2008). Visualization and Level-of-Detail Control for Multi-Dimensional Bioactive Chemical Data. *2008 12th International Conference Information Visualisation*,

11-16. https://doi.org/10.1109/IV.2008.19

Yang, M. S., & Hussain, I. (2023). Unsupervised multi-view K-means clustering algorithm. *IEEE Access*, *11*, 13574-13593. https://doi.org/10.1109/ACCESS.2023.3243133

Zhang, F. (2008). The application of visualization technology on e-commerce data mining. *2008 Second International Symposium on Intelligent Information Technology Application*, *2*, 563-566. https://doi.org/10.1109/IITA.2008.18

Zhu, N., Marais, J., Bétaille, D., & Berbineau, M. (2018). GNSS position integrity in urban environments: A review of literature. *IEEE Transactions on Intelligent Transportation Systems*, *19*(9), 2762-2778. https://doi.org/10.1109/TITS.2017.2766768

Zidan, J., Adegoke, E. I., Kampert, E., Birrell, S. A., Ford, C. R., & Higgins, M. D. (2020). GNSS vulnerabilities and existing solutions: A review of the literature. *IEEE Access*, *9*, 153960-153976. https://doi.org/10.1109/ACCESS.2020.2973759